A Survey on Fault Diagnosis in Wireless Sensor Networks

Zeyu Zhang, *Student Member, IEEE*, Amjad Mehmood, Lei Shu, *Senior Member, IEEE,* Zhiqiang Huo, *Student Member, IEEE,* Yu Zhang and Mithun Mukherjee, *Member, IEEE*

***Abstract*—Wireless sensor networks (WSNs) often consist of hundreds of sensor nodes that may be deployed in relatively harsh and complex environments. In views of hardware cost, sensor nodes always adopt relatively cheap chips, which makes these nodes become error-prone or faulty in the course of their operation. Natural factors and electromagnetic interference could also influence the performance of the WSNs. When sensor nodes become faulty, they may have died which means they cannot communicate with other members in the wireless network, they may be still alive but produce incorrect data, they may be unstable jumping between normal state and faulty state. To improve data quality, shorten response time, strengthen network security, and prolong network lifespan, many studies have focused on fault diagnosis. This survey paper classifies fault diagnosis methods in recent five years into three categories based on decision centers and key attributes of employed algorithms: centralized approaches, distributed approaches, and hybrid approaches. As all these studies have specific goals and limitations, this paper tries to compare them, lists their merits and limits, and propose potential research directions based on established methods and theories.**

***Index Terms*—Wireless sensor networks (WSNs), Industrial wireless sensor network (IWSN), Fault diagnosis, Reliability, Lifetime**

I. Introduction

A. Applications of WSNs

Wireless sensor networks (WSNs) comprise large numbers of sensor nodes and one or several sink node (SN) also called base station (BS). From the perspective of quantity, the number of sensor nodes occupies the largest proportion of components [1].These wireless nodes often contain several sensors and act as ”nerve endings” to apperceive and monitor the physical environment, e.g., the natural environment or a man-made environment. Except for the sensor part, a typical wireless sensor node consists of the following components: (a) microcontroller module, (b) transceiver module, (c) power source module, and (d) additional module, e.g., mobilizer, actuator, etc. These nodes utilize radio channels to communicate with each other and share their information, which can be forwarded to a SN or BS directly or by multi-hop relays.

WSNs have been widely applied to various applications [2]–[5]. One early example, possibly the first application of WSNs, is the air-delivered seismic intrusion detector (ADSID) system [6]. This system was used by United States Air Force (USAF) in the Vietnam War to detect the Vietnamese transport troops as it was known the harsh environment of the tropical rainforest made the US military suffer. In this system, each node was equipped with a sensitive seismometer that was planted along the Ho Chi Minh Trail to detect vibrations from moving vehicles. The sensed data were regularly transmitted from each node directly to an airplane, over a channel with a unique frequency. The system was used to facilitate the dispatch of bombers to strike targets, usually troops moving along the trail. The application of this system greatly decreased American casualties and, in addition, seriously damaged the supply lines of Vietnamese army. In modern military, the applications of WSNs become more extensive, such as self-healing land mines (SHLM) [7], aerostat acoustic payload for transient detection (AAP) [8], soldier detection and tracking (SDT) [9], early attack reaction sensor (EARS) [10], sniper detection and localization (SDL) [11] and so on.

Beyond early military applications, some researchers applied WSNs to agricultural production [12]. The relevant application for crop protection was designed to divert animal intrusions in crop fields since crop damage by animals is one of the major threats to maximizing crop yield [13]. Apart from pest control, WSNs are in use with different agricultural services like irrigation [14], fertilization [15], greenhouse [16] and so on. WSNs can also be applied to health monitoring of the human body, especially for hospital patients or the elderly [17].

Industry is another important orientation and Industrial Wireless Sensor Networks (IWSNs) can be classified into three groups [18]:

- *Environment sensing*, currently represents the widest group of WSN applications [19]. This category is further subdivided into pollution, hazardous environment, and security sensing. Pollution sensing is directed at air, water, gas, and noise pollution, while hazardous environment sensing include fire, flood, landslide, debris flow, and gas leakage sensing. Meanwhile, security sensing presents security issues to be fixed which arises in markets with other competing providers and products, where IWSNs are used for monitoring barrier areas and points of interest.
- *Condition monitoring*, covers equipment, structure, and workers' status monitoring [20]. The status of industrial equipment may worsen over time, which makes it necessary for WSNs to monitor machines' working deviation from optimal situations. WSNs can also be used to equip workers or those working in dangerous situations to decrease casualties and damage, such as for firemen, miners, and etc. These people always face potential dangers in the performance of their duties.
- *Process automation*, refers to the use of WSNs to monitor and control important automated processes and to make the processes more reliable [21].

Z. Zhang is with NAU-Lincoln Joint Research Center of Intelligent Engineering, Nanjing Agricultural University, China, 210000, email: zeyuzhang@outlook.com.

A. Mehmood is with Institute of Information Technology, Kohat University of Science & Technology, Kohat, Pakistan, email: dramjad.mehmood@ieee.org.

L. Shu is with NAU-Lincoln Joint Research Center of Intelligent Engineering, Nanjing Agricultural University, China, 210000 and with School of Engineering in College of Science, University of Lincoln, UK, INB3235, Engineering Hub, Brayford Pool, Lincoln, LN6 7TS, United Kingdom, email: lei.shu@ieee.org. L. Shu is the corresponding author.

Z. Huo, Y. Zhang are with School of Engineering in College of Science, University of Lincoln, UK, INB3235, Engineering Hub, Brayford Pool, Lincoln, LN6 7TS, United Kingdom, email: {zhuo, yzhang}@lincoln.ac.uk.

M. Mukherjee is with Guangdong Provincial Key Lab of Petrochemical Equipment Fault Diagnosis, Guangdong University of Petrochemical Technology, China, 525000, email: mithun.mukherjee@outlook.com.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

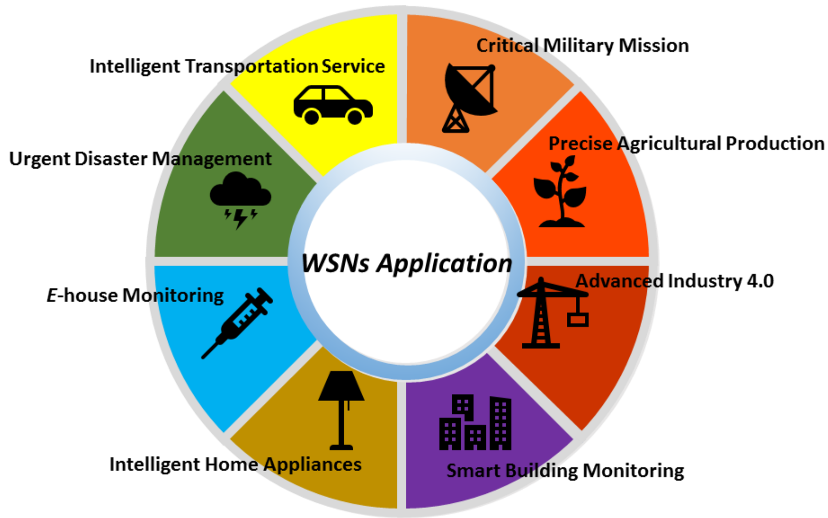

Fig. 1. Overview of WSN applications

In general, WSNs have been mainly used in eight areas, as shown in Fig. 1.

B. Challenges of Applying WSNs

Despite the wide application of WSNs, technical bottlenecks still exist among them. The three main challenges for WSN applications are the following [22].

1) Reliability: A significant parameter for assessing WSN reliability is the measure of the amount of data received correctly at the SN or BS [23]. Of course, some applications relating to personal security require very high reliability. For example, WSNs employed in battlefield surveillance [24] demand high reliability, and a small mistake could cause heavy casualties. In contrast, agricultural WSNs may require a lower network reliability. Reliability is always a function of the cost of sensor nodes. Consequently, how to properly price a sensor required to deal with real-word scenarios becomes very difficult. In addition, the environment or place where these nodes are located could interfere with the reliability of the network. For example, the industrial environment may be rife with network-compromising factors, such as complex fields, higher or lower temperature, electromagnetic waves generated by machines or wires, chemical substances, and so on. At the same time, multiple wireless HART networks usually coexist in a real industrial environment with such sensors, and could cause performance degradation due to communication interference among networks [25]. Such factors could reduce the reliability of networks and cause relevant errors in the data transport process.

2) Real-time performance: Evaluating the average time a data packet takes to travel from a sensor node to the sink node or base station is a common way of measuring real-time performance [26]. In general, military applications require much quicker response speeds than any other type. Real-time performance is mainly dictated by several key factors, including retransmission time, congestion, MAC delay, etc. Because of recent advances in CPU performance, the primary time consumption derives from communication and message congestion. There are many reasons for time delays in the communication phase. Climatic conditions, such as wind, rain, snow, humidity, temperature, and even solar radiation, can have a negative effect on communication. Beyond these factors, other reasons for impaired network performance may not be attributable to nature. Many sensor nodes are deployed in industrial plants in order to monitor or sense various aspects of the environmental or mechanical environments. The rotation of machines, moving objects, and even chemical reagents can also influence the performance of a network. Hence, the influencing factors are complex and difficult to estimate.

3) Safety: A major consideration in many fields, especially the military, is safety [27]. Since WSNs are open networks and the energy and hardware requirements are limited, developers cannot implement highly complicated algorithms or routing protocols to guarantee the safety of a network. It is believed that a strong receiver deployed in the area of a WSN can easily decode the data packets of most WSNs, and even deliver the wrong data to the base station in order to force the system to make an incorrect judgment. The consideration of safety may be unnecessary in most areas except in military applications. In modern warfare, WSNs have been widely used in battlefield monitoring, cooperative operations, executing precise strikes, etc. Because this kind of network is responsible for maintaining connections in the physical world, the safety and accuracy of the required data transmission is very significant.

C. Motivations for fault diagnosis in WSNs

According to the demands of most applications, sensor nodes in WSNs are always expected to work together autonomously in an unattended, harsh, and even hostile environments. Consequently, these nodes tend to be faulty or go dead over time. If the faulty nodes are not tracked and handled correctly in time, they will inevitably lead to data unreliability, affect network bandwidth, cause sectional route congestion, and reduce a network's lifetime. The motivations of fault diagnosis therefore include the following.

1) Increase data reliability: Sensor nodes become faulty and unreliable for different reasons, e.g., hardware and software failure, environment effects, malicious attacks that drive nodes to produce faulty data, etc. These latter data will be forwarded to a SN or BS and decrease the accuracy of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

judgments by the BS. In the area of automation control, node failure can result in large numbers of casualties and property losses, e.g., in the case of networks used to detect poisonous gases in chemical plants. In this scenario, if the network fails to detect the gas leakage in time, worker safety cannot be guaranteed. A specific example is the space shuttle Challenger disaster, an event that undoubtedly increased vigilance regarding sensor integrity. Because one of the sensor nodes failed in the network of one of the shuttle Discovery's external solid fuel tanks, Discovery's scheduled launch was canceled [28]. Fault diagnosis can determine whether the sensed data are normal or faulty, and can eliminate faulty data and decrease the side effects on the SN.

2) Make bandwidth utilization efficient: Bandwidth refers to the rate of data transfer, measured in bits per second. As a wireless sensor network is a resource-constrained network, it is difficult to broaden the bandwidth according to nodes' needs. Faulty data inevitably occupy network bandwidth. The scenario is worse for multimedia sensor networks because there is a huge magnitude of bandwidth required for the network for multimedia data transmission. Furthermore, the algorithms for multimedia sensors require more complex hardware and computation power for processing, making computation energy dissipation equivalent to that of communication energy [29], [30].

3) Prolong network lifetime: The network lifetime relates directly to nodes' energy [31]. Nodes start dying because they have consumed most of their energy performing various network operations and data communication. Low battery power results in poor connectivity between nodes, and when this occurs, the nodes will no longer be the part of the network and instead cause the network to be partitioned. Fault diagnosis can decrease faulty data transmission, which is capable of reducing a network's energy consumption effectively. The dead nodes can also lead to wasted energy consumption. If a dead node was the routing node, it can cause routing loops.

In fact, it is very challenging to develop a protocol that fulfills all features of the aforementioned categories of fault diagnosis in sensor nodes [32], [33]. By considering the importance and challenges of fault diagnosis, we decided to survey existing papers, published from 2013 to 2017, and present a high-level view of fault diagnosis in WSNs.

D. Contribution and organization

1) Contribution: The main contributions of this paper are as follows:

- Supplements the newest research results in fault diagnosis in WSNs from 2013 to 2017, as the most important and recent review of this area was by Mahapatro *et al.* [17], with coverage only up to 2012. Another related review of fault detection, by Muhammed *et al.* [34], focuses mainly on fault detection and does not cover fault diagnosis, actually fault diagnosis is a finer job than fault detection, which have been defined in Table I.
- Analyzes some representative papers and points out their shortcomings and limitations.
- Provides insights into the existing papers and suggests some potential research directions, which may facilitate growth in this area.

2) Organization: The remainder of the paper is organized as follows:

- Description of the basic fault types and research progress from 2013 to 2017 (Section II).

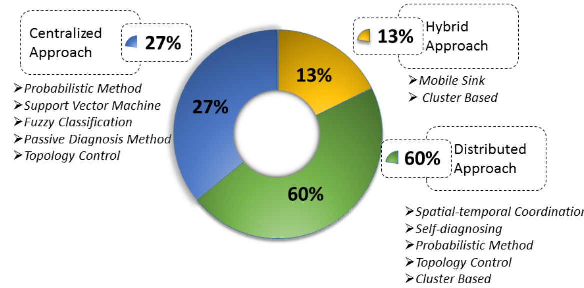

Fig. 2. Fault diagnosis categorization and appropriate approaches

- Summary of the major researcher results and discussions (Section III).
- Introduction of potential future research directions (Section IV).
- Conclusions (Section V).

II. Fault diagnosis approaches in WSNs

A. Approach classification and fundamental terminology

Fault diagnosis approaches have different classification methods and are of three types depending on where the decision of sensor node status is made [35]–[37]:

1) Centralized approach: A geographically or logically centralized node, e.g., central controller or manager, the sink node, takes responsibility for fault management of the overall network.

2) Distributed approach: Every sensor node is able to make decisions at certain levels; the decision center is transferred from the sink node to a common node.

3) Hybrid approach: Between the centralized and distributed approaches, in which both the sink and common nodes have the right to decide the status of nodes.

In order to quickly understand the current state of the literature, we present the most important fault diagnosis terms in Fig. 2 .

B. Fault types

This section contains the common definitions of faults, classification methods, and concrete manifestations to help readers gain a basic understanding of fault types.

1) Fault definitions:

- A fault is an unexpected change or malfunction in a system, although it may not lead to physical failure or breakdown [46].
- Unless ground truth is known or given by something with high confidence, the term fault can only refer to a deviation from the expected model of the phenomenon. A data fault is data reported by a sensor that is inconsistent with the phenomenon of interest's true behavior [40].

2) Fault classifications: There are different ways of classifying fault types found in literature. Generally, faults can be mainly divided into two categories, as shown in Fig. 3:

- *Hard faults*: a sensor node is not capable of communicating with the rest of the network.
- *Soft faults*: a sensor node continues to operate and communicate with altered behavior, e.g., produces faulty data, cannot act as a sable routing node.

Hard faults are also called *permanent faults*. They result from the failures of some hardware modules [47]:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

TABLE I
FAULT DIAGNOSIS TERMINOLOGY

Term	Definition
Active diagnosis	Diagnosis that is attained by active and continuous monitoring of the current state, using the normal state for reference.
Faulty values	This type of faults occurs normally in networks due to processing strategies [38]. A node may generate fault data due to a hardware problem, or might measure or receive faulty values.
Communication failures	This type of failure occurs due to environmental conditions, hardware problems, etc. [39]. When the communication between two nodes is interrupted during a distributed computation.
Tampered nodes	Nodes in WSNs may suffer from malicious activities [40]. An attacker could reprogram the sensor node after taking it over, thus making it follow the attacker's instructions next.
Passive diagnosis	Diagnosis that is attained by passive and intermittent monitoring of the inconsistency of the current state, using the normal state for reference.
Failure detection rate	The rate of detecting faulty nodes from existing faults.
Transient faults	Temporary malfunctions of the computing unit that cause an incorrect result to be computed.
Isolation	Determines which node(s) is(are) faulty [36].
Permanent faults	Faults that are continuous and stable in time, and produce errors when fully exercised.
Intermittent faults	Faults that are repeated occurrences of transient faults. They sometimes cause a faulty sensor node to behave in a fault-free manner, and occur during normal system operations. Thus, they are a highly important class of failure within WSNs.
Fault diagnosis	Consists of (1) detection, (2) isolation, (3) identification, and (4) recovery [41].
Total latency	(Recv Time(n)-Sent Time(n)).
Average latency	(Total Latency / Total Packets Received).
Packet delivery ratio (P-DR%)	Measures the ratio of total packets received against total packets sent.
Fault tolerance	Consists of (1) prevention, (2) isolation, (3) identification, and (4) recovery [42].
Average PDR	(Total Packets Received / Total Packets Sent) * 100.
Energy consumption	Measures energy dissipated by a node while sending packets to the BS.
Prevention	Maintains network connectivity, and provides redundant links/nodes when required [42].
Detection accuracy (DA)	Total number of defected fault nodes/total number of faulty nodes [43].
Fault management	Consists of (1) fault detection, (2) fault diagnosis, and (3) discovering faulty sensor nodes. These techniques repair and resolve faults and failures at any time [43].
Time complexity	The elapsed time between inception and the end of the diagnosis session, often defined as diagnosis latency.
Network lifetime	The operational time of the network during which it performs the assigned or dedicated task(s).
Message complexity	The total number of messages exchanged by nodes during the execution of an algorithm
Detection latency	Maximum time required to detect all faulty sensor nodes present in the network, often considered time complexity.
False alarm rate	The probability of fault-free sensors to be diagnosed as faulty [44].
Fault management	Consists of monitoring network behavior, recognizing the occurrence of any faults, and identifying their type or origin so that an efficient reaction or response can be offered.
Fault detection	Detecting whether there are fault nodes present in the network [36].
Fault hypotheses	Introduces a latent fault, analyzes its outcome, and presents some useful method(s) of dealing with it.
Fault identification	Specifies the type of fault that occurred [36].
Fault recovery	Estimates the output of the faulty nodes [36].
Model-based	The status of a sensor node is decided by a model [13].
Offline detection	Performed by a wired network or delay-tolerant applications [13].
Online detection	Real-time detection [13].
Model-less	Deciding the status of a sensor node without a model.
Classification accuracy (CA)	Number of nodes classified in a particular class/total number of nodes in that class. Classes include the (i) permanent fault class, (ii) intermittent fault class, and (iii) fault-free class [45].

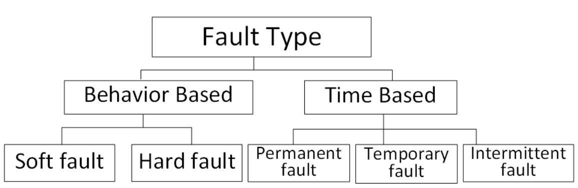

Fig. 3. Broad categorization of fault types

- *Communication module faults* or *transceiver module faults*
- *Battery depletion*
- *Out of communication range of entire mobile network*

Soft faults are always *temporary* or *intermittent*, which means nodes with soft faults act arbitrarily and are difficult to predict and detect [48], [49]:

- *Byzantine*, a node behaves arbitrarily or maliciously.
- *Omission*, a failure by omission is determined by a service sporadically not responding to requests.
- *Timing*, timing failure occurs when a node responds to a request out of the time interval, which is always in the situation that demands higher real-time performance.

C. Centralized approach

In centralized approaches, one centralized sensor node, always a SN or BS, is responsible for performing fault management. The statuses of the other nodes are decided

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

by this centralized node, which possesses high computational power, abundant memory size, and persistent energy supply. Most times, the centralized node receives information from the rest of nodes proactively or passively. By analyzing the information, the centralized node can confirm the statuses of the other nodes. In terms of papers published in recent years, the centralized approaches can be classified into the following methodologies.

1) Probabilistic method: In the probabilistic method, fault diagnosis is considered a pattern-classification problem. Many classification algorithms are applied to this problem, e.g., the naïve Bayes classification algorithm and the maximum posterior probability hypothesis.

Bill *et al.* [50] proposed the centralized naïve Bayes detector to classify sensor nodes by analyzing the end-to-end transmission time collected at the sink. This approach is described in Table II.

As with the CNDB algorithm, Tang *et al.* [51] modeled the network as a graph using an extended algorithm known as a neighborhood hidden conditional random field (NHCRF), as shown in Table III. The NHCRF judges a faulty sensor node in the network by collecting its signal strength, frequency, and signal delay. It subsequently relaxes the independent assumptions that help determine nonlocal dependencies among states and observations. Thus, the status between sensor nodes and transmission paths can be determined. Furthermore, because of inclusion between state dependencies, performance evaluations show that NHCRF is very effective and efficient at fault diagnosis under different sizes and traffic loads. Furthermore, Dhal *et al.* [52] proposed an approach that regards a classification problem as a maximum posterior probability hypothesis testing problem.

Discussion: Both of the algorithms described in the aforementioned two papers cannot know the number of fault nodes in advance, which limits their application. Meanwhile, as a fault detection method, the probabilistic method cannot distinguish various faults in WSNs.

2) Support vector machine (SVM): The SVM is one type of supervised learning model with associated learning algorithms that analyze data used for classification and regression analysis in machine learning. Yu *et al.* [44] proposed a new direction of fault node diagnosis. Their algorithm tried to reduce fault information in order to decrease diagnosis time. As we know, fault diagnosis has significant communications overhead, calculation complexity, and large energy consumption. This paper claims to use rough set (RS) theory to filter out less important data and build a new simple dataset that is used to train the SVM. Therefore, RS-SVM fault diagnosis is done using the aforementioned methods. Furthermore, RS-SVM can effectively and diagnose and detect faulty network nodes more accurately than other methods, as shown in Table IV.

Discussion: Sensor node faults can be classified into many types, and it is difficult for SVM to perform this task.

3) Fuzzy classification: This approach is also one of the common machine learning approaches. Compared to the probabilistic method, observed values do not have a necessary relation to a certain status. One certain status is always decided by several observed values, and each value has its own weight. Chanak *et al.* [53] demonstrated a fuzzy rules-based faulty node classification and management scheme (FNCM) for the detection of physical and environmental conditions, e.g., road monitoring, smart home automation, and livestock management. It distinguishes itself from existing approaches in four ways. First, it uses an efficient data routing algorithm for the recovery and reusability of faulty nodes. Second, it overcomes the problem of uncertainty. Third, it assigns work to a node per its hardware capabilities and status. Finally, its management of nodes not only helps to achieve an efficient routing scheme, but also increases overall network performance.

Discussion: Fuzzy classification requires a better understanding of WSNs and internal relationships. Meanwhile, the faulty types have to be confirmed first.

4) Passive diagnosis method: The above methods can be classified as proactive approaches, which often make energy depletion quicker and reduce network lifespan, as these algorithms place extra communications overhead on networks. In order to overcome this drawback, Xiaohang *et al.* [54] proposed a passive anomaly detection model based on the autoregressive (AR) model and non-parametric tests (the Kuiper and K-S tests). Like a rough set, the AR model acts as a linear prediction filter to pre-whiten the test data. The diagnosed data and data travel times, are randomly picked up and transmitted to the sink node. The diagnosed data of normal conditions are generated and used for training the non-parametric test. If one routing node fails, as a sensor node employs *ad hoc* on-demand distance vector routing (AODV) routing, the new route will be longer than the original one. In other words, in this paper, anomalous conditions, including two parameters, i.e., traffic conditions and number of faulty nodes, will make their signals deviate from the normal ones. At this stage, then, K-S and Kuiper tests are used to indicate the difference.

Discussion: Just like the SVM approach, this method cannot differentiate between specific fault types and determine the number of fault nodes.

5) Topology control: Occasional errors occurring in the network inevitably cause a change of topology. This type of fault can be classified as an intermittent or transient fault. Christopher *et al.* [55] proposed the SEDEL Sensor nEtwork DEfect Localization (SEDEL) method to deal with this problem. The authors model the routing topology of each processing stage as a graph, or a tree. A WSN operator, a centralized node or sink node, is used to store the routing topology using an echo-based topology-discovery algorithm [56]. A graph-mining approach, i.e., a frequent subgraphs mining approach, is used to detect the frequent subgraphs database, as shown in Fig. 4. These subgraphs can be used to generate a table that contains class labels and edges. In the next step, information gain (InfoGain) for all nodes is calculated from that table. Finally, by using the output of InfoGain, a node's status is decided based on its ranking. Performance evaluation showed that this technique narrows the defected node's location in the routing table to, at most, two neighboring nodes. It also helps mark occasional errors that are usually difficult to track or detect.

Discussion: The method cannot deal with the topology change when two or more defective nodes exist at the same time.

6) Drawbacks of centralized diagnosis:

- Centralized fault diagnosis algorithms cannot be applied to large-scale WSNs, as each sensor node requires multiple hops to communicate with a sink node or BS, which depletes the energy of routing nodes quickly, especially the nodes located around the central node.
- Large-scale WSNs can also lead to significant diagnostic delay, which means that the status of remote node could change while the sink node is making a judgment.
- The sink node is responsible for diagnosing the statuses of all the other sensor nodes. However, if the sink node

TABLE II
CENTRALIZED NAÏVE BAYES DETECTOR

STEP 1	Preparatory phase	1. Confirm the features, i.e., end-to-end packet transmission time. 2. Obtain the training samples.
STEP 2	Training phase	1. Estimate the conditional probability of MLE. 2.Build naïve Bayes classifier. 3. Estimate marginal probability.
STEP3	Testing phase	1. Compare normal and faulty conditional probabilities of the mostly observed delay. 2. Directed at the faulty network, further differentiate the reasons, i.e. network congestion or faulty node.

TABLE III
NHCRF METHOD FOR FAULT DETECTION

STEP 1	Modeling phase	1. Collect historical data 2. Features and labels. 3. Train the parameter
STEP 2	Monitoring phase	1. Collect newly arriving data. 2. Features. 3. Compute the states of nodes and the label for the entire input.
STEP3	Results phase	1. The states of nodes, fault nodes diagnosis result. 2. The label of the input, faulty scenes diagnosis result.

TABLE IV
ALGORITHM STRUCTURE OF WSN FAULT DIAGNOSIS BASED ON RS-SVM INFORMATION FUSION

STEP 1	Sample formation and data preprocessing	1. Determine fault type. 2. Preprocess the original signals. 3. Filter less important data with rough set.
STEP 2	Simulation and training for SVM	1. Train and select the best kernel function to determine the SVM parameters.
STEP3	Training for SVM with training reduction	1. Reconstitute the test sample set with minimum condition attributes and corresponding initial data as input; the output is the final result of diagnosis

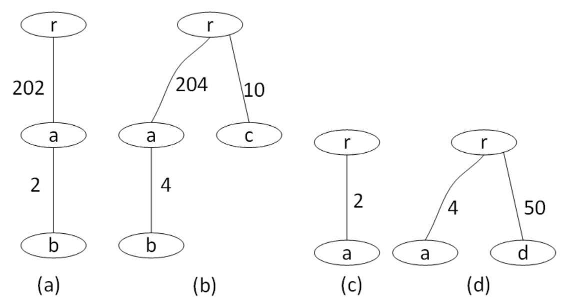

Fig. 4. Example of routing-tree database: owing to a software error, node a always adds 200 to the value of node b ((a) and (b)), but follows the protocol correctly in (c) and (d).

is faulty, the diagnosing process cannot be completed properly.

- Real-time performance cannon be guaranteed, as the status of every sensor node must decided by the sink node. Hence, when network size grows, real-time performance worsens.
- The diagnosis latency is very high, as the sink node requires a global view of the entire network before it can make a decision on the status of every sensor node.

D. Distributed approach

Unlike the centralized approach, each sensor node in the model-less or distributed approach makes decisions about their health status by gathering and analyzing diagnostic response results from neighboring nodes. Then, they update the BS accordingly. Therefore, the model-less approach transfers a little information to the BS, which helps prolong network lifetime. It further reduces much traffic overhead, and minimizes the end-to-end delay over the network. There are many recent techniques described in the literature that follow distributed approaches for fault detection and diagnosis.

1) Spatial-temporal coordination: In this kind of approach, diagnosis methods depend on spatial and temporal coordination. In terms of spatial coordination, one sensor node, e.g., node s_i, is used to monitor the local temperature, and another sensor node, s_j, is the neighboring node, which means it is in the transmission radius of s_j. Both nodes have similar data. In terms of temporal coordination, one node, if it is fault free, has relatively stable data over a period of time.

Miao *et al.* [57] demonstrated agnostic diagnosis to discover silent failures in WSNs. This is a sink-based technique that collects data from all sensor nodes in the network. This technique is different from other techniques in the following ways: (i) it does not consider predefined rules; it relies on *a priori* knowledge as little as possible, and it can be applied to a large number of applications in WSNs; (ii) it generates a correlation graph that can efficiently characterize correlations between metrics and can describe the latent status inside a node; and (iii) it demonstrates an agnostic diagnosis (AD) algorithm, an online lightweight failure detection approach, and checks its effectiveness through a 330-GreenOrb-node deployment. The effectiveness of this algorithm was demonstrated through studies of different cases and statistical analysis. Furthermore, since it is a sink-based technique, there is a delay between fault time and fault detection time.

In fault detection, there is a trade-off between detection accuracy and detection latency. More tests or operations on the status of one node are certain to improve detection accuracy, but can also lead to superior detection latency or detection delay. During this period, the status of a node may change. Arunanshu *et al.* [39] proposed a method based on multi-objective swarm optimization to solve this problem. This fault detection method still depends on neighboring nodes, so it can be classified as spatial-temporal coordination.

Hamdana *et al.* [58] illustrated a test and diagnostic technique for WSN applications. This technique deals with two fault classes. In the first, it considers the node faults and link failures (predefined faults). In the second, it tracks failures occurring at system dysfunctionalities or application levels. The proposed technique covers the following: (i) helping the protocol diagnose network faults; (ii) assessing the impact of faults on the entire network; (iii) helping improve the fault detection rate by using some predefined faults; and (iv) after any changes or code upgrades, validating the application according to operating conditions. It also helps improve monitoring at different levels without imposing significant overhead. Lastly, this work provides valuable information to the administrator to facilitate handling the problem quickly or even temporarily

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

ignoring it.

Lo *et al.* [36] proposed a distributed model-based nonlinear sensor fault diagnosis in WSNs. This model suggests that model-based communication consumes more energy due to greater communication between nodes than in distributed approaches [59]–[62]. This study used sensor input and output values to determine non-linearity faults. According to the study, nodes suffering from a non-linearity fault have normal and abnormal working regions. Furthermore, nodes provide correct measurements when the true signal falls in the normal region, and distorted measurements when the true signal resides in the abnormal region. This protocol works locally between every pair of sensor nodes, therefore saving a significant amount of energy compared to centralized fault diagnosis strategies.

Discussion: This kind of method depends largely on the node degree and the status of neighboring nodes, i.e., the diagnostic accuracy of one node would decrease, while most of its neighboring nodes are faulty or have few neighboring nodes.

2) Self-diagnosing: In this type of approach, sensor nodes are required to compare their sensor data with that of their neighbors. Sensor node statuses are determined by the neighbors. These algorithms can work properly at the early stage of deployment, as most sensor nodes are normal and their judgments are correct. As time goes by, the performance of algorithms degrade, especially the common mode failures (CMFs), which are impossible for comparative methods to detect.

In this approach, to reduce the effect of neighboring nodes' data, a sensor node is capable of detecting its own status. Babaie *et al.* [43] suggested a new self-diagnosing approach. This approach reduces the effect of neighboring nodes and uses Petri nets and a correlation graph to analyze the behavior of sensor nodes. By using Petri nets, which are actually flow charts, sensor nodes are capable of detecting different kinds of faults, i.e., permanent faults and transient faults. The correlation graph is used to diagnose the failure of inner links between sensor components.

Panda *et al.* [63] proposed a distributed self fault diagnosis (DSFD) algorithm to solve the fault diagnosis problem of large-scale WSNs, and this approach can diagnose both hard and soft faults. It divides fault diagnosis into two phases, i.e., the initialization phase and self-diagnosis phase. In the first phase, actually a data-collecting phase, writers assume that all sensor nodes are fault free. After this phase, every sensor node can have a local view of its neighboring nodes and their sensed data. Another important assumption is that the sensed data of every node follows normal distribution. In the self-diagnosis phase, when node s_i cannot receive information from a neighboring node s_j, s_j is considered to be hard faulty. If s_i can receive data from the neighboring node s_j, s_j will perform a 3σ test to identify whether s_j is soft faulty. For normal distribution, the probability of data remaining in (μ-3σ, μ+3σ) is 0.9974. A 3σ test uses the normal distribution assumption to test whether the sensed data is in the range of larger probability. If not, s_j is soft faulty, as shown in Table V.

Distributed systems always suffer different kinds of soft faults. A Byzantine fault is one of them, i.e., a faulty node may exhibit arbitrary behavior, e.g., a faulty node may corrupt its local state and send arbitrary messages. A Byzantine fault is intermittent and difficult to predict. Meenakshi *et al.* [33] proposed a fault detection technique based on hypothesis testing, which is like that described in [63] by the same authors, which uses the idea of small probability. In this self-detectable distributed fault detection algorithm, each sensor node collects data from its neighbors and diagnoses itself by using the Neyman-Pearson test from hypothesis testing theory.

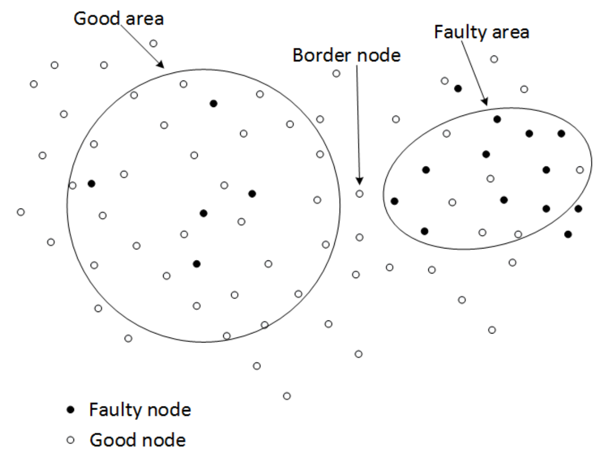

Fig. 5. Relative position of faulty area, good area, and border node.

Discussion: A self-diagnosing approach demands stringent assumptions and still depends on the first-hand data from neighboring nodes. Thus, future research should relax the assumptions and decrease the dependence on neighboring nodes.

3) Probabilistic method: In the centralized approach, we introduced the probabilistic method to diagnose node status. As a sink node has no energy limitation and enough computation power, it can build a classifier to diagnose node status. However, a centralized probabilistic method suffers high diagnosis latency. Yuan *et al.* [32] proposed a distributed Bayesian algorithm (DBA), as shown in Table VI. This method uses a border node to adjust the fault probability calculated by a neighboring node, which efficiently decreases the negative effect caused by faulty neighboring nodes, as presented in Fig. 5

Chafiq Titouna *et al.* [64] also presented a fault detection scheme (FDS) for WSNs. Their method used probabilistic classifiers employing the formalism of Bayesian networks. This method represents the network in the form of a directed acyclic graph that shows a probability distribution. Each node is represented by a random variable X_i, and the edge between two nodes shows a probabilistic dependency of a child. The network structure illustrates that each X_i from its parent is conditionally independent from its non-descendants. According to these assumptions, a conditional probability table is associated, illustrating that each X_i distribution assigns any possible values to its parents. A Bayesian network is simply a Bayesian classifier used for task classification.

In this scheme, a class variable is represented by C, and X_i represents each node feature. Hence, in order to calculate the probability $P(C=c_k | X=x)$ for each possible class c_k, Bayes theorem is given in Eq. (1). It is not often possible to achieve $P(C=c_k | X=x)$ without making independent assumptions. The most important assumption in the naive Bayesian classifier is that each feature X_i is independent of each and every available feature in the class variable C, as shown in Eq. (2). The proposed technique determines the conditional probability of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

TABLE V
DSFD

STEP 1	Initialization phase	1. Two assumptions, all sensor nodes are fault-free, sensed data follows normal distribution. 2. Every sensor node has a local view of its neighboring nodes and sensed data.
STEP 2	Self-diagnosis phase	1. The node is hard faulty or not by its neighboring nodes. 2. The soft fault is decided by a 3σ test.

TABLE VI
DBA. P_i AND P_j REFERS TO PRIOR FAULT PROBABILITY OF NODE S_i, S_j RESPECTIVELY. F_{ij} REFERS TO THE FLAG OF NODES' STATUS, IF S_i AND S_j ARE IN DIFFERENT STATUS, F_{ij}=0, OTHERWISE, F_{ij}=1

STEP 1	Calculating the probability	1. Every node compares its sensed reading with neighbors to get f_{ij} 2. Calculate p_{si}.
STEP 2	Adjusting the probability	1. Adjusting the fault probability p_{si} by exploiting the border node. 2. Deciding whether one node is border node or node. 3. Border node sends a message to its neighbors to obtain their confidence c.
STEP 3	Judging status of node	1. If the fault probability is higher than the probability threshold, the node will be considered a faulty node.

sensor node N_i, which gives the remaining energy level EL_i^t, and the sensed data SD_i^t at time t is shown in Eqs. (3) and (4), respectively, using Bayes' rule. We have

$$P(X = x|C = c) = \prod P(Xi = xi|C = c), \tag{1}$$

$$P_i^t(N_i|SD_i^t) = \frac{P(SD_i^t|N_i)P(N_i)}{P(SD_i^t)}, \tag{2}$$

$$P_i^t(N_i|EL_i^t) = \frac{P(EL_i^t|N_i)P(N_i)}{P(EL_i^t)}. \tag{3}$$

After determining the result, the probability joint (PJ) is sent to its CH. Then, a decision is made on the basis of similarity existing among the PJs of all neighboring nodes belonging to the same cluster. Performance evaluations illustrated that this FDS outperformed FDWSN.

Discussion: The method still relies on neighboring nodes. In [32], the authors tried to adjust the fault probability using the border node, but the detection accuracy will decrease when time elapses, as more nodes become faulty, and it is difficult to differentiate good areas and faulty areas.

4) Topology control: Sensor nodes with limited energy will be dead when their battery power is exhausted. If these nodes are routing nodes, they could affect network connectivity. Compared to the energy scarcity in the latter stage, most energy is wasted in the first stage, when most nodes are deployed closely and can communicate with all the nodes in their transmission range, and it is unnecessary and wasteful. In order to prolong the lifetime of a network, Deniz *et al.* [42] proposed an adaptive energy-aware and distrusted fault-tolerant topological-control algorithm (an adaptive disjoint path vector, or ADPV, algorithm), as schematically depicted in 6. The protocol works in two phases: initialization and restoration. In the initialization phase, the ADPV finds all alternative paths based on a set-picking method pre-existing in the network. The restoration phase initiates whenever k-vertex connectivity with the stationary supernode is broken. To restore connectivity, the ADPV uses the calculated alternative paths and readjusts the nodes' transmission ranges accordingly. The ADPV is distributed in nature, and simulation results have illustrated that it prolongs the lifespan of heterogeneous nodes connected to the supernode. It also guarantees network connectivity durability, ranging from 5% to 95%, against node failures. Moreover, in cases of 75% and 90% node failure, the network remains connected to the supernode through three or two vertexes, respectively.

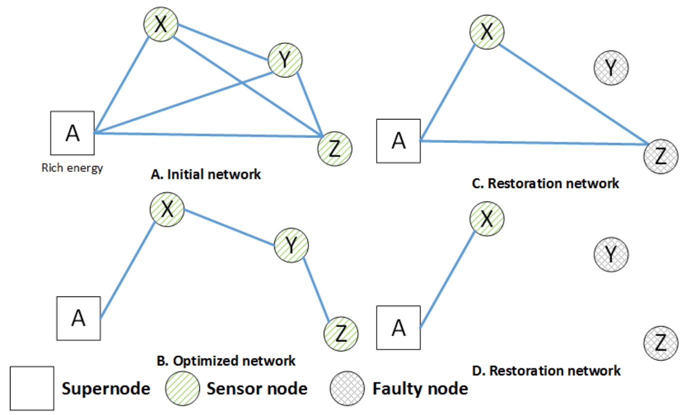

Fig. 6. Adaptive disjoint path vector algorithm: Figures A and B are the initial stages of node deployment. A common node will adjust its transmission range to connect to the supernode with the smallest communication overhead. Over time, some common nodes die owing to battery depletion or hard faults. In Figs. C and D, original routing is broken, and the rest of the nodes readjust the transmission ranges.

Discussion: This method mainly focuses on making the most of limited energy. According to the connectivity, which is related to the change of node status, nodes adjust their transmission ranges. A supernode, which can also be regarded as a cluster head, is responsible for collecting data from the common nodes.

5) Cluster Based: A cluster head is a kind of supernode with rich energy and abundant computation capability, the characteristics of which are adaptive for performing fault diagnosis in WSNs. Mehdi *et al.* [65] proposed a fault-tolerant service (FTS) based on a hierarchical network. This service can be divided into three steps:

- Fault detection: this step can be divided into two types, cluster head (CH) fault detection and cluster member (CM) fault detection, as shown in Fig. 7. A CH fault is detected by a spare cluster head, the neighboring CHs, and the CMs. Fault detection of a CM is accomplished by the CHs. Both CHs and CMs have to send heartbeat, summary, and update messages periodically to their corresponding nodes. If these corresponding nodes receive none of these messages, the CHs or CMs are considered to be faulty and the process advances to the next step.
- Fault diagnosis: the FTS uses time redundancy to detect transient faults around both CMs and CHs. If a fault cannot be affirmed to be transient fault, it is a permanent fault.
- Fault recovery: CH fault recovery is a replacement of the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

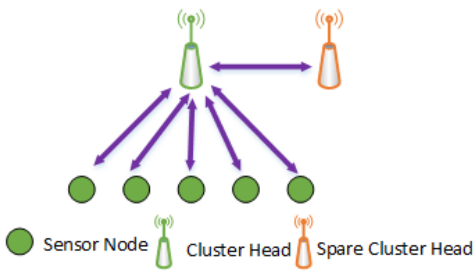

Fig. 7. Fault-tolerant service fault detection steps.

spare cluster head. CM recovery is removal of the faulty CMs from the routing table.

Discussion: Compared to other distributed algorithms, a cluster-based method for fault diagnosis in WSNs is simple and sufficient. The detection accuracy of this method is higher than that of other distributed methods at the expense of hardware costs. This fault diagnosis method still must rely on the judgment of neighboring nodes; thus, the algorithm's performance will worsen over time.

6) Drawbacks of distributed diagnosis: A distributed fault diagnosis approach gives the decision back to node level. Compared to a centralized approach, a distributed method can be applied to large-scale networks. However, it still has some drawbacks demanding urgent solutions.

- Algorithm simplification; that is, common sensor nodes in WSNs are always equipped with limited energy and computation capacity. Thus, most machine learning algorithms cannot be applied to distributed approaches.
- Most distributed approaches rely on the performance of neighboring nodes, but over time most nodes become faulty, which inevitably has a negative effect on detection accuracy.
- In the diagnosis process, the status of a common node is decided by its sensed data with neighboring nodes whether exceeding the threshold or not. In most cases, the threshold is decided by experience, which cannot be applied in uncharted territory.

E. Hybrid approach

Based on decision center, a hybrid approach has two decision components, one in the sink node and one in the common nodes. This approach originated from the acknowledgement of two main drawbacks in centralized and distributed approaches. A centralized approach cannot be applied in large-scale networks, and has a relatively higher diagnosis latency. The main problem in a distributed method is keeping the detection accuracy high. In order to solve these problems, a hybrid approach was proposed. Up to now, the basic thought of a hybrid approach has been to add extra equipment, e.g., a mobilizer, to achieve diagnosis reliability, robustness, energy efficiency, and minimization of traffic overhead.

1) Mobile sink: In order to overcome the limitations in both distributed and centralized approaches, and due to the improvisational nature of WSNs, the lack of insight into internal running status, and, in particular, since network structure can frequently change due to link failure, Chanak *et al.* [66] presented a mobile sink-based distributed fault detection scheme, which identifies the health status of each software and hardware component separately. In this algorithm, the mobile detector starts its fault diagnosis from the BS. As it explores each deployed node, it obtains its health status. It then uploads the information from all nodes in the network. It completes its operation by returning to the BS. This information helps the administrator recover and reuse faulty sensor nodes. It also helps maintain reliability, and improves the lifespan of the network. Experiments concluded that this scheme outperforms existing fault detection techniques because single-hop communication for detection is followed.

Lastly, Zahhad *et al.* [67] illustrated a mobile sink-based adaptive immune energy-efficient clustering protocol (M-SIEEP) to deal with elevating the energy hole problem by using a controlled mobile sink, which is based on the fact that the nodes near the BS die earlier than those far away. The MSIEEP helps ensure node connectivity with the BS. It also divides the network into small regions so the nodes require a smaller range of radio transmissions than when in a full area, as shown in Fig. 8. Moreover, this adaptive protocol helps find sojourn locations for the mobile sink, and considers the energy dissipation during communication and from overhead control packets in the network. The MSIEEP also decides an optimal number of CHs, and their locations. Overall, this protocol improves the overall lifetime, network connectivity, packet drop rate, and security of the network. Simulation results indicated that this technique is more energy efficient and reliable than existing techniques.

Discussion: The mobile sink based method has many advantages over centralized or distributed methods. The main problem with a mobile sink based method lies in the path planning of mobile sink nodes. The two papers cited illustrate two different ways this is done. In [66], the optimal diagnostic hub points come from the center of the triangle formation between the deployed sensor nodes. In [67], the sensor field is divided into small regions, e.g., rectangles, and the center of the rectangle is chosen as the sojourn location. The path planning of mobile sink nodes is greatly related to the algorithm's performance. The shortest path obviously decreases the diagnosis delay and promotes data reliability.

2) Cluster-based method: In the distributed method, Mehdi *et al.* [65] proposed a cluster-based method, where the network is divided into three levels, i.e., common node level, cluster head level, and sink node level. Zafar *et al.* [68] suggested an analogous hybrid fault detection mechanism that performs fault detection. The nature of this method ... The diagnostic entities in the network can be divided into three parts:

- Diagnostic agent (DA), which periodically monitors sensor node (SN) processes.
- Local cluster head, a local diagnosis center that performs the diagnostic processes in a cluster.
- Error-specific cluster head, which stores a certain distributed error database used to deal with certain types of errors.

Discussion: The nature of the cluster-based method uses hardware redundancy to improve fault detection accuracy. This method can effectively decrease the complexity of the algorithm, but inevitably increases the cost of WSNs.

3) Drawbacks of hybrid diagnosis :

- In the mobile-sink-based method, path planning determines the algorithmic performance. As one of the hybrid methods discussed in this paper, the mobile-sink-based algorithm has greater detection delay. However, the selection of hub points in the path and path planning according to these points is a typical non-deterministic polynomial complete problem.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

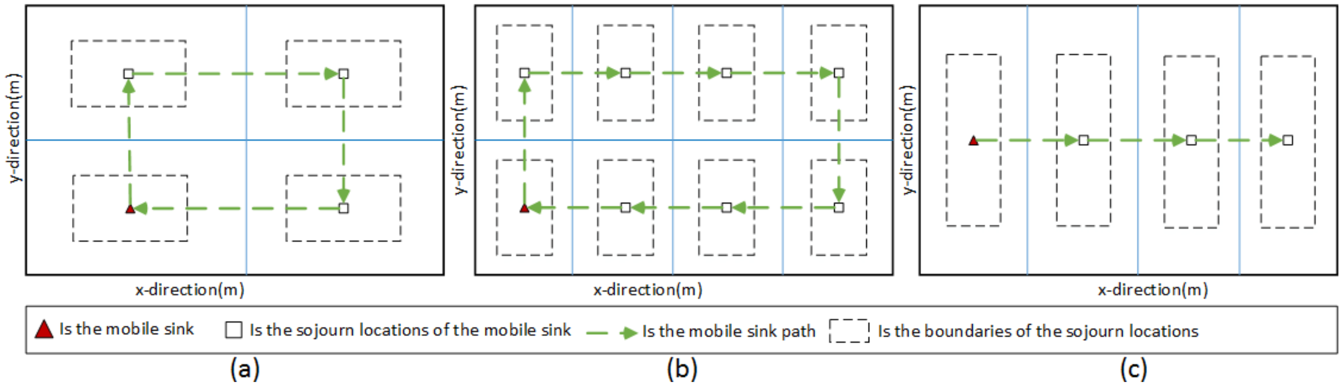

Fig. 8. Sink mobility patterns. (a) Four regions, rectangular path pattern. (b) Eight regions, rectangular path pattern. (c) Four regions, line path pattern.

- The cluster-based method is like a local centralized method. Supernodes are added into the networks to perform local diagnosis tasks, which increases the construction cost of networks.

III. Summary and Insights of Main Contributions

A. Summary of main contributions

This section summarizes the main conclusions of papers published from 2013 to 2017, as shown in Table VII.

B. Insights of major research

We present the insights of major research, to illustrate potential future research directions, in Table VIII.

1) Mobile WSNs: Mobile WSNs (MWSNs) consist of mobile sensors or sink nodes in the networks [76]. The advantages of MWSNs over static WSNs are better energy efficiency, improved coverage, enhanced target tracking, and superior channel capacity. MWSNs have a much more dynamic topology compared to static WSNs. The proposed methods for fault diagnosis in static WSNs always perform poorly in MWSNs.

2) Industrial WSNs: Industrial wireless sensor networks (IWSNs) are used for controlling and monitoring various industrial tasks [22]. WSNs have great advantages over wired networks, e.g., no cables, cost reduction, and ease of installation and repair. Aside from the advantages, IWSNs are faced with unique challenges caused by industrial control systems, e.g., reliability, real-time communication, and robustness, which establishes new requirements for WSN fault diagnosis.

3) Wireless multimedia sensor networks (WMSNs): Currently, most deployed WSNs are limited to collect scalar data, e.g., temperature, humility, location, light intensity and pressure. However, there are also useful multimedia data, e.g., video, voice, and image, in the areas of medical care, traffic monitoring, and smarthomes. WMSNs consist of sensor nodes that can collect video, voice and image data. Compared to traditional WSNs, WMSNs require higher energy consumption, bandwidth, and quality of service.

4) Underwater WSNs: Wireless underwater sensor networks (WUSNs) refer to WSNs deployed underwater, e.g., under lakes, rivers, and oceans. Since sensor nodes are usually deployed in the deep sea and typically powered by batteries, they can hardly be recharged or are not worth replacing [77]. Unlike terrestrial WSNs, where the locations of the sensor nodes can be determined by global positioning system technology, underwater sensor nodes can be localized through limited communication with anchor or reference nodes. Owing to the fluctuation of water, the deployment of WUSNs involves a three-dimensional (3D) environment. Sensor nodes deployed underwater are always used to monitor the environment and detect certain events. Special underwater environments could influence the reliability of sensed data and communication. All of these new scenes call for higher requirements in fault diagnosis. So far, few researchers have focused on the fault diagnosis of WUSNs.

5) 3D-based WSNs: Sensor nodes are often deployed in 3D areas, not just underwater, such as military sensing devices deployed in a nonplanar battlefield, a sensor network floating in the air for tracking chemical plumes, or a fire monitoring network in a mountainous forest [78]. According to the papers on fault diagnosis in WSNs published from 2013 to 2017, although two-dimensional (2D) WSNs have been widely explored, the networks in which sensor nodes have three dimensions have been less thoroughly researched. A 3D network has different traits than a 2D network, e.g., the topology of a 3D network is more complex than that of a 2D network, which challenges the existing diagnosis methods relying on space correlation in 2D networks.

6) Software-defined-network (SDN)-based WSNs: An SDN is a centralized network structure in which every computation is completed in the controller rather than in the sensors themselves, and all information is exchanged only through switches [79]. The pivotal technology of SDNs is called OpenFlow, which can separate the computational unit and transmission unit of a single device. In general, an SDN has the following advantages over traditional WSNs:

- The use and control of an SDN can be programmed, which provides considerably more configuration flexibility.
- An SDN deceases the hardware cost of network since the computational unit of a node is removed.
- An SDN contributes to visualization of the network and effectively combines network computing and storage resources.
- The controller is responsible for computing, and for decreasing both the error ratio and potential exterior interference.

The main faults of an SDN originate in data communication and storage. As common nodes have no computational unit, nearly all distributed algorithms cannot be employed in an SDN. Moreover, centralized methods also cannot run in an SDN directly, which requires further research in the future.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

TABLE VII
ANALYSIS OF EXISTING PROTOCOLS WITH RESPECT TO DIFFERENT FAULT DIAGNOSIS PARAMETERS

Author	Year	Diagnosis network			Diagnosis view		Persistence of fault			Fault type		Diagnosis approach
		Distributed	Centralized	Hybrid	Local	Global	Permanent	Intermittent	Transient	Hard	Soft	
Shahram *et al.* [43]	2013	✓			✓	✓	✓		✓	✓	✓	Self-diagnosing
Miao *et al.* [57]	2013	✓				✓		✓	✓		✓	Spatial-temporal coordination
Kulla *et al.* [69]	2013		✓		✓			✓	✓	✓	✓	Probabilistic
Alessandra *et al.* [70]	2013	✓			✓			✓	✓		✓	Probabilistic
Dima *et al.* [58]	2013	✓			✓		✓	✓	✓	✓	✓	Spatial coordination
Banerjee *et al.* [71]	2014	✓			✓	✓	✓	✓	✓	✓	✓	Spatial-temporal coordination
Bill *et al.* [50]	2014		✓			✓		✓	✓		✓	Probabilistic
Arunanshu *et al.* [39]	2014	✓			✓		✓	✓		✓	✓	Spatial-temporal coordination
Manmath *et al.* [45]	2014	✓			✓		✓	✓		✓	✓	Spatial-temporal coordination
Mehdi *et al.* [65]	2014	✓			✓		✓	✓	✓	✓	✓	Cluster-based
Yu *et al.* [44]	2014		✓			✓	✓	✓	✓	✓	✓	SVM
M. Panda *et al.* [63]	2014	✓			✓		✓	✓	✓	✓	✓	Self-diagnosing
Yuan *et al.* [32]	2015	✓			✓			✓			✓	Probabilistic
Zafar *et al.* [35]	2015			✓	✓	✓		✓	✓		✓	Cluster-based
Dhal *et al.* [52]	2015		✓			✓		✓	✓		✓	Topology control
Gong *et al.* [72]	2015		✓			✓		✓	✓		✓	Topology control
Meenakshi *et al.* [33]	2015	✓			✓			✓	✓		✓	Probabilistic
Lo *et al.* [36]	2015			✓	✓	✓		✓	✓		✓	Spatial coordination
Chafiq *et al.* [64]	2015	✓			✓	✓		✓	✓		✓	Probabilistic
Jin *et al.* [54]	2015		✓		✓	✓		✓	✓		✓	Model-based
Mohammed *et al.* [67]	2015			✓		✓	✓	✓	✓	✓	✓	Mobile sink-based
Chanak *et al.* [53]	2016	✓			✓		✓		✓	✓	✓	Spatial coordination
Christopher *et al.* [55]	2016		✓			✓		✓	✓		✓	Topology control
Panigrahi *et al.* [30]	2016	✓			✓			✓	✓		✓	Spatial coordination
Zhen *et al.* [73]	2016	✓				✓		✓	✓		✓	Cluster-based
Hongsheng *et al.* [29]	2016	✓			✓		✓	✓	✓	✓	✓	Spatial coordination
Zhang *et al.* [74]	2016	✓			✓	✓	✓	✓	✓	✓	✓	Spatial coordination
Tang *et al.* [51]	2016	✓			✓	✓		✓	✓		✓	Probabilistic method
Chanak *et al.* [66]	2016			✓	✓	✓	✓	✓	✓	✓	✓	Mobile sink-based
Sujie *et al.* [75]	2017		✓		✓		✓	✓	✓	✓	✓	Spatial-temporal coordination

7) Energy harvesting for WSNs: Energy shortages are the primary bottleneck in WSN applications and for fault diagnosis. Most sensor nodes are equipped with a battery that has limited energy that is quickly depleted. This drawback limits the ability of nodes to undertake execution of relatively complex fault diagnosis algorithms. Energy harvesting or power harvesting is the process by which sensor nodes both obtain energy from the external world, e.g., solar energy, wind energy, and thermal energy, and store the energy. This orientation involves the energy of WSNs, and sensor nodes must be equipped with extra hardware, which demands improved algorithms in fault diagnosis.

8) Inductive charging for WSNs: Inductive, or wireless, charging employs an electromagnetic field to transmit energy from one object to another. As in energy harvesting for WSNs, inductive charging also involves the energy-related problem of WSNs. Different from energy harvesting, the energy of inductive charging always comes from artificial electromagnetic waves. This is the same situation as in the medium that nodes use to communicate with each other. Furthermore, this orientation also demands improved algorithms for fault diagnosis.

9) Heterogeneous Wireless Sensor Networks (HWSNs): HWSNs consist of nodes with different capabilities in terms of hardware and protocols. Most fault diagnosis methods only consider homogeneous wireless networks, in which sensor nodes are the same in terms of both hardware and protocol; thus, the fault diagnosis method cannot applied to HWSNs directly.

10) Duty-cycle-based WSNs: To prolong the lifetime of the WSNs, one common approach is to dynamically schedule the sensors' awake/sleep cycles, i.e., duty cycle or sleep scheduling [80]; the cited work describes a connected k-neighborhood (CKN)-based approach (k refers to the least-awake neighbors). In existing fault diagnosis algorithms, sensor nodes in WSNs are active until dead; thus, duty-cycle-based WSNs must consider related algorithmic improvement.

IV. CONCLUSIONS

Since WSNs have limited resources and are usually deployed in inaccessible, uncontrolled, and autonomous environments, each node in the network must be monitored to avoid adverse effects of faulty nodes on normal network operations.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2794519, IEEE Access

TABLE VIII
INSIGHT OF MAJOR RESEARCHES: 1. MOBILE WSNS; 2. INDUSTRIAL WSNS; 3. MULTIMEDIA WSNS; 4. UNDERWATER WSNS; 5. 3D ENVIRONMENT-BASED WSNS; 6. SDN-BASED WSNS; 7. ENERGY HARVESTING FOR WSNS; 8. WIRELESS CHARGING FOR WSNS; 9. HETEROGENEOUS-NETWORK-BASED WSNS; 10. DUTY-CYCLE-BASED WSNS

Author	Year	1	2	3	4	5	6	7	8	9	10
Shahram *et al.* [43]	2013	×	×	×	×	×	×	×	×	×	×
Miao *et al.* [57]	2013	×	×	×	×	×	×	×	×	×	×
Kulla *et al.* [69]	2013	×	×	×	×	×	×	×	×	×	×
Alessandra *et al.* [70]	2013	×	×	×	×	×	×	×	×	×	×
Dima *et al.* [58]	2013	×	×	×	×	×	×	×	×	×	×
Banerjee *et al.* [71]	2014	×	×	×	×	×	×	×	×	×	×
Bill *et al.* [50]	2014	×	×	×	×	×	×	×	×	×	×
Arunanshu *et al.* [39]	2014	×	×	×	×	×	×	×	×	×	×
Manmath *et al.* [45]	2014	×	×	×	×	×	×	×	×	×	×
Mehdi *et al.* [65]	2014	×	×	×	×	×	×	×	×	×	×
Yu *et al.* [44]	2014	×	×	×	×	×	×	×	×	×	×
M. Panda *et al.* [63]	2014	×	×	×	×	×	×	×	×	×	×
Yuan *et al.* [32]	2015	×	×	×	×	×	×	×	×	×	×
Zafar *et al.* [35]	2015	×	×	×	×	×	×	×	×	×	×
Dhal *et al.* [52]	2015	×	×	×	×	×	×	×	×	×	×
Gong *et al.* [72]	2015	×	×	×	×	×	×	×	×	×	×
Meenakshi *et al.* [33]	2015	×	×	×	×	×	×	×	×	×	×
Lo *et al.* [36]	2015	×	×	×	×	×	×	×	×	×	×
Chafiq *et al.* [64]	2015	×	×	×	×	×	×	×	×	×	×
Jin *et al.* [54]	2015	×	×	×	×	×	×	×	×	×	×
Mohammed *et al.* [67]	2015	✓	×	×	×	×	×	×	×	×	×
Chanak *et al.* [53]	2016	×	×	×	×	×	×	×	×	×	×
Christopher *et al.* [55]	2016	×	×	×	×	×	×	×	×	×	×
Panigrahi *et al.* [30]	2016	×	×	×	×	×	×	×	×	×	×
Zhen *et al.* [73]	2016	×	✓	×	×	×	×	×	×	×	×
Hongsheng *et al.* [29]	2016	×	×	×	×	×	×	×	×	×	×
Zhang *et al.* [74]	2016	×	×	×	×	×	×	×	×	×	×
Tang *et al.* [51]	2016	×	×	×	×	×	×	×	×	×	×
Chanak *et al.* [66]	2016	✓	×	×	×	×	×	×	×	×	×
Sujie *et al.* [75]	2017	×	✓	×	×	×	×	×	×	×	×

Low-cost sensor nodes often become error prone and unreliable due to hardware, software, and/or other imperfections manifesting as "glitches." Consequently, fault diagnosis is required to identify, detect, isolate, reuse, or let the fault-free sensor work to address faulty events. This allows the network to be operational even in the presence of faults.

Fault diagnosis can be observed at either side of the network, such as at the BS (centralized), node sides (distributed), or a combination of both (hybrid). Hybrid networks have a larger picture of the whole network compared to that in the node-based approach, and therefore decisions can be made from a relatively broader perspective. The node side avoids traffic overhead and delay, which increases the overall lifetime of the network. As a result, the hybrid approach achieves the advantages of the other approaches while avoiding their disadvantages. Thus, by using this approach, a better fault diagnosis protocol or algorithm can be proposed. Significant work has been done on sorting out the issues of reliability, robustness, and lifetime in WSNs [45], [47], [72]. This survey provides a broader picture of current promising techniques for fault detection and diagnosis. It also elaborates their strong and weak points. We believe that this survey will be helpful in proposing more robust, reliable, scalable, real-time, mobile, energy-efficient and intelligent protocols in the near future.

ACKNOWLEDGEMENTS

This work was partially supported by International and Hong Kong, Macao, and Taiwan collaborative innovation platforms, by major international cooperation projects of colleges in Guangdong Province (Grant No. 2015KGJHZ026), and by the Natural Science Foundation of Guangdong Province (Grant No.2016A030307029), Maoming Engineering Research Center of Industrial Internet of Things (No.517018).

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubaramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

[2] A. N. Harun, R. Ahmad, and N. Mohamed, "Wsn application in led plant factory using continuous lighting (cl) method," in *2015 IEEE Conference on Open Systems (ICOS)*, Aug 2015, pp. 56–61.

[3] A. E. Kouche, L. Al-Awami, H. Hassanein, and K. Obaia, "Wsn application in the harsh industrial environment of the oil sands," in *2011 7th International Wireless Communications and Mobile Computing Conference*, July 2011, pp. 613–618.

[4] X. Lu, I. H. Kim, A. Xhafa, and J. Zhou, "Wsn for machine area network applications," in *2016 46th European Solid-State Device Research Conference (ESSDERC)*, Sept 2016, pp. 23–28.

[5] M. P. ?uri?i?, Z. Tafa, G. Dimi?, and V. Milutinovi?, "A survey of military applications of wireless sensor networks," in *2012 Mediterranean Conference on Embedded Computing (MECO)*, June 2012, pp. 196–199.

[6] S. Singhal, A. K. Gankotiya, and S. Agarwal, "An investigation of wireless sensor network: A distributed approach in smart environment," in *Advanced Computing & Communication Technologies*. IEEE, 2012, pp. 522–529.

[7] W. M. Merrill, L. Cirod, B. Schiffer, D. Mclntire, G. Rava, K. Sohrabi, F. Newberg, J. Elson, and W. Kaiser, "Dynamic networking and smart sensing enable next-generation landmines," *IEEE Pervasive Computing*, vol. 3, no. 4, pp. 84–90, Oct 2004.

[8] H.-E. de Bree and J. W. Wind, "The acoustic vector sensor: a versatile battlefield acoustics sensor," 2011. [Online]. Available: http://dx.doi.org/10.1117/12.884681

[9] B. Rippin, "Pearls of wisdom wireless networks of miniaturized sensors," 2012. [Online]. Available: http://dx.doi.org/10.1117/12.918116

[10] J. Chang, W. Mendyk, L. Thier, P. Yun, A. LaRow, S. Shaw, and W. Schoenborn, "Early attack reaction sensor (ears), a man-wearable gunshot detection system," 2006. [Online]. Available: http://dx.doi.org/10.1117/12.665978

[11] S.Hengy, P.Hamery, S. Mezzo, and P.Duffner, "Networked localization of sniper shots using acoustics," 2011. [Online]. Available: http://dx.doi.org/10.1117/12.883906

[12] A. ur Rehman, A. Z. Abbasi, N. Islam, and Z. A. Shaikh, "A review of wireless sensors and networks' applications in agriculture," *Computer Standards & Interfaces*, vol. 36, no. 2, pp. 263–270, 2014.

[13] V. Bapat, P. Kale, V. Shinde, N. Deshpande, and A. Shaligram, "Wsn application for crop protection to divert animal intrusions in the agri-

cultural land," *Computers and Electronics in Agriculture*, vol. 133, pp. 88–96, 2017.

[14] G. H. E. L. de Lima, L. C. e Silva, and P. F. R. Neto, "Wsn as a tool for supporting agriculture in the precision irrigation," in *2010 Sixth International Conference on Networking and Services*, March 2010, pp. 137–142.

[15] S. Khan, "Wireless sensor network based water well management system for precision agriculture," in *2016 26th International Telecommunication Networks and Applications Conference (ITNAC)*, Dec 2016, pp. 44–46.

[16] K. P. Ferentinos, N. Katsoulas, A. Tzounis, T. Bartzanas, and C. Kittas, "Wireless sensor networks for greenhouse climate and plant condition assessment," *Biosystems Engineering*, vol. 153, no. Supplement C, pp. 70 – 81, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1537511016302847

[17] A. Mahapatro and P. M. Khilar, "Fault diagnosis in wireless sensor networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2000–2026, 2013.

[18] H. A. Salam and B. M. Khan, "Iwsn - standards, challenges and future," *IEEE Potentials*, vol. 35, no. 2, pp. 9–16, March 2016.

[19] K. Mikhaylov, J. Tervonen, J. Heikkil?, and J. K?ns?koski, "Wireless sensor networks in industrial environment: Real-life evaluation results," in *2012 2nd Baltic Congress on Future Internet Communications*, April 2012, pp. 1–7.

[20] J. Neuzil, O. Kreibich, and R. Smid, "A distributed fault detection system based on iwsn for machine condition monitoring," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1118–1123, May 2014.

[21] G. Alderisi, G. Patti, O. Mirabella, and L. L. Bello, "Simulative assessments of the ieee 802.15.4e dsme and tsch in realistic process automation scenarios," in *2015 IEEE 13th International Conference on Industrial Informatics (INDIN)*, July 2015, pp. 948–955.

[22] H. A. Salam and B. M. Khan, "IWSN–Standards, Challenges and Future," *IEEE Potentials*, vol. 35, no. 2, pp. 9–16, 2016.

[23] F. Huang, Z. Jiang, S. Zhang, and S. Gao, "Reliability evaluation of wireless sensor networks using logistic regression," in *2010 International Conference on Communications and Mobile Computing*, vol. 3, April 2010, pp. 334–338.

[24] I. Bekmezci and F. Alagoz, "Energy efficient, delay sensitive, fault tolerant wireless sensor network for military monitoring," in *Sensors Applications Symposium, 2008. SAS*. IEEE, 2008, pp. 172–177.

[25] X. Jin, F. Kong, L. Kong, W. Liu, and P. Zeng, "Reliability and temporality optimization for multiple coexisting wirelesshart networks in industrial environments," *IEEE Transactions on Industrial Electronics*, vol. PP, no. 2, p. online, 2017.

[26] Y. Li, C. S. Chen, Y. Q. Song, Z. Wang, and Y. Sun, "Enhancing real-time delivery in wireless sensor networks with two-hop information," *IEEE Transactions on Industrial Informatics*, vol. 5, no. 2, pp. 113–122, May 2009.

[27] S. Roy and M. J. Nene, "A security framework for military application on infrastructure based wireless sensor network," in *2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*. IEEE, 2015, pp. 369–376.

[28] Markets and Markets, *Liftoff of Our 2015 Assessment of NASAs Major Space Projects!*, 08 2016. [Online]. Available: https://blog.gao.gov/tag/national-aeronautics-and-space-administration/

[29] H. Xu, R. Zhang, C. Lin, and Y. Ma, "Novel approach of fault diagnosis in wireless sensor networks node based on rough set and neural network model," *International Journal of Future Generation Communication and Networking*, vol. 9, no. 4, pp. 1–16, 2016.

[30] T. Panigrahi, M. Panda, and G. Panda, "Fault tolerant distributed estimation in wireless sensor networks," *Journal of Network and Computer Applications*, vol. 69, pp. 27–39, 2016.

[31] F. Kerasiotis, A. Prayati, C. Antonopoulos, C. Koulamas, and G. Papadopoulos, "Battery lifetime prediction model for a wsn platform," in *2010 Fourth International Conference on Sensor Technologies and Applications*, July 2010, pp. 525–530.

[32] H. Yuan, X. Zhao, and L. Yu, "A distributed bayesian algorithm for data fault detection in wireless sensor networks," in *2015 International Conference on Information Networking (ICOIN)*. IEEE, 2015, pp. 63–68.

[33] M. Panda and P. Khilar, "Distributed byzantine fault detection technique in wireless sensor networks based on hypothesis testing," *Computers & Electrical Engineering*, vol. 48, pp. 270–285, 2015.

[34] T. Muhammded and R. A. Shaikh, "An analysis of fault detection strategies in wireless sensor networks," *Journel of Network and Computer Applications*, vol. online, pp. 267–287, 2016.

[35] A. Zafar, B. Wajid, and B. A. Akram, "A hybrid fault diagnosis architecture for wireless sensor networks," in *International Conference on Open Source Systems & Technologies (ICOSST)*. IEEE, 2015, pp. 7–15.

[36] C. Lo, J. P. Lynch, and M. Liu, "Distributed model-based nonlinear sensor fault diagnosis in wireless sensor networks," *Mechanical Systems and Signal Processing*, vol. 66, pp. 470–484, 2016.

[37] M. Yu, H. Mokhtar, and M. Merabti, "Fault management in wireless sensor networks," *IEEE Wireless Communications*, vol. 14, no. 6, pp. 1536–1284, 2007.

[38] T. Tosic, N. Thomos, and P. Frossard, "Distributed sensor failure detection in sensor networks," *arXiv preprint arXiv:1109.5636*, 2011.

[39] A. Mahapatro and A. K. Panda, "Choice of detection parameters on fault detection in wireless sensor networks: A multiobjective optimization approach," *Wireless Personal Communications*, vol. 78, no. 1, pp. 649–669, 2014.

[40] K. Ni, N. Ramanathan, M. N. H. Chehade, L. Balzano, S. Nair, and *et al.*, "Sensor network data fault types," *ACM Transactions on Sensor Networks (TOSN)*, vol. 5, no. 3, pp. 1–29, 2009.

[41] *Research and Markets: Industrial Wireless Sensor Network Sensor & Technology Market Trends & Forecasts to 2013-2020 for the Oil & Gas, Energy & Power, Automotive, & Food & Beverage Industries*, 08 2016. [Online]. Available: http://www.businesswire.com/news/home/20150108005601/en/Research-Markets-Industrial-Wireless-Sensor-Network-Sensor visited on 2016-09-21

[42] F. Deniz, H. Bagci, I. Korpeoglu, and A. Yazıcı, "An adaptive, energy-aware and distributed fault-tolerant topology-control algorithm for heterogeneous wireless sensor networks," *Ad Hoc Networks*, vol. 44, pp. 104–117, 2016.

[43] S. Babaie, A. Khosrohosseini, and A. Khadem-Zadeh, "A new self-diagnosing approach based on petri nets and correlation graphs for fault management in wireless sensor networks," *Journal of Systems Architecture*, vol. 59, no. 8, pp. 582–600, 2013.

[44] C.-B. Yu, J.-J. Hu, R. Li, S.-H. Deng, and R.-M. Yang, "Node fault diagnosis in wsn based on rs and svm," in *Wireless Communication and Sensor Network (WCSN), 2014 International Conference on*. IEEE, 2014, pp. 153–156.

[45] M. N. Sahoo and P. M. Khilar, "Diagnosis of wireless sensor networks in presence of permanent and intermittent faults," *Wireless Personal Communications*, vol. 78, no. 2, pp. 1571–1591, 2014.

[46] C. Hajiyev and F. Caliskan, *Fault Diagnosis and Reconfiguration in Flight Control Systems*, ser. 1572–1272. Springer US, 2003, vol. 2.

[47] S. Chessa and P. Santi, "Crash faults identification in wireless sensor networks," *Computer Communications*, vol. 25, no. 14, pp. 1273–1282, 2002.

[48] M. Panda and P. M. Khilar, "Distributed soft fault detection algorithm in wireless sensor networks using statistical test," in *2012 2nd IEEE International Conference on Parallel, Distributed and Grid Computing*, Dec 2012, pp. 195–198.

[49] M. Panda and P. Khilar, "Distributed byzantine fault detection technique in wireless sensor networks based on hypothesis testing," *Computers & Electrical Engineering*, vol. 48, no. Supplement C, pp. 270 – 285, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0045790615002293

[50] B. C. Lau, E. W. Ma, and T. W. Chow, "Probabilistic fault detector for wireless sensor network," *Expert Systems with Applications*, vol. 41, no. 8, pp. 3703–3711, 2014.

[51] P. Tang and T. W. Chow, "Wireless sensor-networks conditions monitoring and fault diagnosis using neighborhood hidden conditional random field," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 933–940, 2016.

[52] R. Dhal, J. A. Torres, and S. Roy, "Detecting link failures in complex network processes using remote monitoring," *Physica A: Statistical Mechanics and its Applications*, vol. 437, pp. 36–54, 2015.

[53] P. Chanak and I. Banerjee, "Fuzzy rule-based faulty node classification and management scheme for large scale wireless sensor networks," *Expert Systems with Applications*, vol. 45, pp. 307–321, 2016.

[54] X. Jin, T. W. Chow, Y. Sun, J. Shan, and B. C. Lau, "Kuiper test and autoregressive model-based approach for wireless sensor network fault diagnosis," *Wireless Networks*, vol. 21, no. 3, pp. 829–839, 2015.

[55] C. Oßner, E. Buchmann, and K. Böhm, "Identifying defective nodes in wireless sensor networks," *Distributed and Parallel Databases*, pp. 1–20, 2016.

[56] M. Bestehorn, Z. Benenson, E. Buchmann, and *et al.*, "Query dissemination in sensor networks—predicting reachability and energy consumption," *Ad Hoc & Sensor Wireless Networks*, vol. 9, pp. 85–107, 2008.

[57] X. Miao, K. Liu, Y. He, D. Papadias, Q. Ma, and Y. Liu, "Agnostic diagnosis: Discovering silent failures in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 6067–6075, 2013.

[58] D. Hamdan, I. Parissis, A. Hijazi, B. El Hassan *et al.*, "Test and diagnosis of wireless sensor networks applications," in *Computer and Information Technology (WCCIT), 2013 World Congress on*. IEEE, 2013, pp. 1–7.

[59] J. Chen, S. Kher, and A. Somani, "Distributed fault detection of wireless sensor networks," in *Proceedings of the 2006 workshop on Dependability issues in wireless ad hoc networks and sensor networks*. ACM, 2006, pp. 65–72.

[60] S. Guo, Z. Zhong, and T. He, "Find: faulty node detection for wireless sensor networks," in *Proceedings of the 7th ACM conference on embedded networked sensor systems*. ACM, 2009, pp. 253–266.

[61] M. Ding, D. Chen, K. Xing, and X. Cheng, "Localized fault-tolerant event boundary detection in sensor networks," in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 2. IEEE, 2005, pp. 902–913.

[62] X. Luo, M. Dong, and Y. Huang, "On distributed fault-tolerant detection in wireless sensor networks," *IEEE Transactions on Computers*, vol. 55, no. 1, pp. 58–70, 2006.

[63] M. Panda and P. M. Khilar, "Distributed self fault diagnosis algorithm for large scale wireless sensor networks using modified three sigma edit test," *Ad Hoc Networks*, vol. 25, pp. 170–184, 2014.

[64] C. Titouna, M. Aliouat, and M. Gueroui, "Fds: fault detection scheme for wireless sensor networks," *Wireless Personal Communications*, vol. 86, no. 2, pp. 549–562, 2016.

[65] M. M. Afsar, "Maximizing the reliability of clustered sensor networks by a fault-tolerant service," in *Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on*. IEEE, 2014, pp. 1–8.

[66] P. Chanak, I. Banerjee, and R. S. Sherratt, "Mobile sink based fault diagnosis scheme for wireless sensor networks," *Journal of Systems and Software*, 2016.

[67] M. Abo-Zahhad, S. M. Ahmed, N. Sabor, and S. Sasaki, "Mobile sink-based adaptive immune energy-efficient clustering protocol for improving the lifetime and stability period of wireless sensor networks," *IEEE Sensors Journal*, vol. 15, no. 8, pp. 4576–4586, 2015.

[68] A. Zafar, B. Wajid, and B. A. Akram, "A hybrid fault diagnosis architecture for wireless sensor networks," in *International Conference on Open Source Systems & Technologies (ICOSST)*. IEEE, 2015, pp. 7–15.

[69] J. Kullaa, "Detection, identification, and quantification of sensor fault in a sensor network," *Mechanical Systems and Signal Processing*, vol. 40, no. 1, pp. 208–221, 2013.

[70] A. De Paola, G. Lo Re, F. Milazzo, and M. Ortolani, "Qos-aware fault detection in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2013, 2013.

[71] I. Banerjee, P. Chanak, H. Rahaman, and T. Samanta, "Effective fault detection and routing scheme for wireless sensor networks," *Computers & Electrical Engineering*, vol. 40, no. 2, pp. 291–306, 2014.

[72] W. Gong, K. Liu, and Y. Liu, "Directional diagnosis for wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 5, pp. 1290–1300, 2015.

[73] Z. Feng, J. Fu, and W. Su, "The gateway anomaly detection and diagnosis in wsns," *in 28th Chinese Control and Decision Conference (CCDC)*, pp. 2401–2406, 2016.

[74] Y. Zhang and X. Yuan, "Fault diagnosis in clustering wsn based on neighbor cooperation," in *2016 Chinese Control and Decision Conference (CCDC)*. IEEE, 2016, pp. 1803–1807.

[75] S. Shao, S. Guo, and X. Qiu, "Distributed fault detection based on credibility and cooperation for wsns in smart grids," *Sensors*, vol. online, pp. 1–14, 2017.

[76] S. A. Munir, B. Ren, and W. Jiao, "Mobile wireless sensor network: Architecture and enabling technologies for ubiquitous computing," in *Advanced Information Networking and Applications Workshops, 2007, AINAW '07. 21st International Conference on*, vol. online. IEEE, May 2007, pp. 1–8.

[77] Z. Zhou, W. Fang, J. Niu, L. Shu, and M. Mukherjee, "Energy-efficient event determination in underwater wsns leveraging practical data prediction," *IEEE Transactions on Industrial Informatics*, vol. online, pp. 1–11, 2017.

[78] C.-H. Ou and K.-F. Ssu, "Sensor position determination with flying anchors in three-dimensional wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 9, pp. 1536–1233, 2008.

[79] Z. Yuan, L. Wang, L. Shu, T. Hara, and Z. Qin, "A balanced energy consumption sleep scheduling algorithm in wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 9, pp. 1536–1233, 2011.

[80] Y. Wang, H. Chen, X. Wu, and L. Shu, "An energy-efficient sdn based sleep scheduling algorithm for wsns," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*. IEEE, May 2011, pp. 1–6.

Zeyu Zhang is a visiting scholar at NAU-Lincoln Joint Research Center of Intelligent Engineering, Nanjing Agricultural University. He received his B.S. from Wuhan Institute Of Technology, Wuhan, China, in 2012, and his M.S. from Wuhan University of Technology in 2015. His research interests are in the field of fault diagnosis in wireless sensor networks.

Amjad Mehmood earned his Ph.D. degree in Wireless Networks from the Kohat University of Science & Technology, Kohat, in 2014. In March 2017, he earned post-doc from Guangdong Provincial Key Laboratory on Petrochemical Equipment Fault Diagnosis at Guangdong University of Petrochemical Technology, Maoming, China under the supervisor of Lei Shu. He is interested to work in the areas of Cyber-Physical Systems, Internet of Things, Connected Vehicles, Wireless Communications and Networking, Optical Communications & Networking, Smart Grid Communications & Networking, and Body Sensor Networks.

Lei Shu (M07-SM15) is currently Lincoln Professor at the University of Lincoln, UK, and a Distinguished Professor at Nanjing Agricultural University, China. He is also the Director of NAU-Lincoln Joint Research Center of Intelligent Engineering, China. His main research field is wireless sensor networks. He has published over 300 papers in related conferences, journals, and books in the area of sensor networks. He had been awarded the Globecom 2010 and ICC 2013 Best Paper awards. He has served as Editor-in-Chief for EAI Endorsed Transactions on Industrial Networks and Intelligent Systems, and as associate editor for *IEEE Systems Journa*l and *IEEE Access*, among others. He has served as co-chair for more than 50 international conferences and workshops, e.g., IWCMC, ICC, ISCC, ICNC, and Chinacom, including roles as Symposium Co-Chair for IWCMC 2012 and ICC 2012; General Co-Chair for Chinacom 2014, Qshine 2015, Collaboratecom 2017, and Mobiquitous 2018; Steering and TPC Chair for InisCom 2015; and TPC member of more than 150 conferences, e.g., ICDCS, DCOSS, MASS, ICC, Globecom, ICCCN, WCNC, and ISCC.

Zhiqiang Huo Zhiqiang Huo is currently working towards his Ph.D. degree at University of Lincoln, UK. He received his Ms. and BS. from China University of Geosciences Beijing, China in 2016 and 2013 respectively. His research interests lie in the ?eld of fault diagnosis of industrial systems, wireless sensor networks and participatory sensing. He had been awarded the INISCOM 2017 Best Paper award. He has served as web chairs in international conferences, such as CollaborateCom 2017, AINIS 2015 and 2016.

Yu Zhang is currently a Senior Lecturer at the School of Engineering, University of Lincoln, Lincoln, UK. She obtained her B.Sc. degree from the School of Aerospace Engineering and Applied Mechanics, Tongji University, Shanghai, China, in 2004. She finished her M.Sc. and Ph.D. degrees from the Department of Civil Engineering, University of Nottingham, Nottingham, UK, in 2005 and 2011, respectively. . Her research interests include equipment fault detection and diagnosis, grey-box system modeling, and development of data analysis and machine learning algorithms. Her recent major projects, including two Innovate UK projects, one international project with Guangdong University of Petrochemical Technology, and one industrial project funded by Siemens, Germany, all focus on the areas of data analysis and machine fault diagnosis.

Mithun Mukherjee received his Ph.D. in electrical engineering from the Indian Institute of Technology, Patna, India in 2015. Currently, he is a Specially Assigned Researcher in the Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis, Guangdong University of Petrochemical Technology, China. He has been serving as the Special Issue Editor of the EAI-endorsed journal *Transactions on Industrial Networks and Intelligent Systems*. His research interests include wireless sensor networks, wireless communications, energy harvesting, and cloud computing.